

Algorithms for Analysing the Temporal Structure of Discourse^{*†}

Janet Hitzeman, Marc Moens and Claire Grover

HCRC Language Technology Group

The University of Edinburgh

2, Buccleuch Place

Edinburgh EH8 9LW, Scotland

J.Hitzeman@ed.ac.uk

Abstract

We describe a method for analysing the temporal structure of a discourse which takes into account the effects of tense, aspect, temporal adverbials and rhetorical structure and which minimises unnecessary ambiguity in the temporal structure. It is part of a discourse grammar implemented in Carpenter's ALE formalism. The method for building up the temporal structure of the discourse combines constraints and preferences: we use constraints to reduce the number of possible structures, exploiting the HPSG type hierarchy and unification for this purpose; and we apply preferences to choose between the remaining options using a temporal centering mechanism. We end by recommending that an underspecified representation of the structure using these techniques be used to avoid generating the temporal/rhetorical structure until higher-level information can be used to disambiguate.

the temporal component were to yield a detailed representation of the temporal structure of the discourse, taking into account the effect of tense, aspect and temporal expressions while at the same time minimising unnecessary ambiguity in the temporal structure. The method combines a constraint-based approach with an approach based on preferences: we exploit the HPSG type hierarchy and unification to arrive at a temporal structure using constraints placed on that structure by tense, aspect, rhetorical structure and temporal expressions, and we use the *temporal centering* preferences described by (Kameyama et al., 1993; Poesio, 1994) to rate the possibilities for temporal structure and choose the best among them.

The starting point for this work was Scha and Polanyi's discourse grammar (Scha & Polanyi 1988; Prüst *et al* 1994). For the implementation we extended the HPSG grammar (Pollard and Sag, 1994) which Gerald Penn and Bob Carpenter first encoded in ALE (Carpenter, 1993). This paper will focus on our temporal processing algorithm, and in particular on our analysis of narrative progression, rhetorical structure, perfects and temporal expressions.

1 Introduction

In this paper we describe a method for analysing the temporal structure of a discourse. This component was implemented as part of a discourse grammar for English. The goals of

2 Constraints on narrative continuations

Probably the best known algorithm for tracking narrative progression is that developed by Kamp (1979), Hinrichs (1981), and Partee (1984), which formalises the observation that an event will occur *just after* a preceding event, while a state will *overlap* with a preceding event. This algorithm gives the correct results in examples such as the following:

- (1) John entered the room. Mary stood up.

^{*}We would like to thank Alex Lascarides and Massimo Poesio for comments on an earlier draft.

[†]This work was supported in part by the European Commission's programme on *Linguistic Research and Engineering* through project LRE-61-062, "Towards a declarative theory of discourse."

- (2) John entered the room. Mary was seated behind the desk.

In (1) the event of Mary's standing is understood to occur just after John enters the room, while in (2) the state in which Mary is seated is understood to overlap with the event of John's entering the room.

However, if there is a rhetorical relationship between two eventualities such as causation, elaboration or enablement, the temporal defaults can be overridden, as in the following examples:

- (3) a. John fell. Mary pushed him.
b. Local builders constructed the Ford St. Bridge. They used 3 tons of bricks.

In (3a) there is a causal relationship between Mary's pushing John and his falling, and the second event is understood to precede the first. In (3b), the second sentence is an elaboration of the first, and they therefore refer to aspects of the same event rather than to two sequential events.

It has been suggested that only world knowledge allows one to detect that the default is being overridden here. For example, Lascarides & Asher (1991) suggest that general knowledge postulates (in the case of (3a): that a pushing can cause a falling) can be invoked to generate the backward movement reading.

The problem for practical systems is twofold: we could assume that in the case of narrative the Kamp/Hinrichs/Partee algorithm is the default, but each time the default is applied we would need to check all our available world knowledge to see whether there isn't a world knowledge postulate which might be overriding this assumption. Clearly this would make the processing of text a very expensive operation.

An alternative is to assume that the temporal ordering between events in two consecutive sentences can be any of the four possibilities (*just-after*, *precede*, *same-event* and *overlap*). But then the resulting temporal structures will be highly ambiguous even in small discourses. And sometimes this ambiguity is unwarranted. Consider:

- (4) Mary stared at John. He gave her back her slice of pizza.

Here, it would appear, only one reading is possible, i.e. the one where John gave Mary her slice of pizza *just after* she stared or started to stare at him. It would be undesirable for the temporal processing mechanism to postulate an ambiguity in this case.

Of course, sometimes it is possible to take advantage of certain cue words which either indicate or constrain the rhetorical relation. For example, in (5) the order of the events is understood to be the reverse of that in (1) due to the cue word *because* which signals a causal relationship between the events:

- (5) John entered the room because Mary stood up.

As Kehler (1994) points out, if forward movement of time is considered a default with consecutive event sentences, then the use of "because" in (5) should cause a temporal clash—whereas it is perfectly felicitous. Temporal expressions such as *at noon* and *the previous Thursday* can have a similar effect: they too can override the default temporal relations and place constraints on tense. In (6), for example, the default interpretation would be that John's being in Detroit overlaps with his being in Boston, but the phrase *the previous Thursday* overrides this, giving the interpretation that John's being in Detroit precedes his being in Boston:

- (6) John was in Boston. The previous Thursday he was in Detroit.

This suggests that the temporal information given by tense acts as a weaker constraint on temporal structure than the information given by temporal adverbials.

The possibilities for rhetorical relations (e.g., whether something is narration, or elaboration, or a causal relation) can be further constrained by aspect. For example, a state can elaborate another state or an event:

- (7) a. Mary was tired. She was exhausted.
b. Mary built a dog house. It was a labour of love.

But an event can only elaborate another event, as in (8):

- (8) a. Mary built a dog house. She used two tons of bricks.
b. Mary was tired/working hard. ?She built a dog house.

Table 1. Possible relations when S_2 expresses a simple past event.

S_1	Relation	Example
past event	just-after S_1 precede S_1 overlap S_1 same-event S_1	Mary pushed John. He fell. John fell. Mary pushed him. NO I assembled the desk myself. The drawers only took me ten minutes.
past activity	just-after S_1 precede S_1 overlap S_1 same-event S_1	Mary stared at John. He gave her back her slice of pizza. NO NO NO
past state	just-after S_1 just-after TF_1 precede S_1 precede TF_1 overlap S_1 overlap TF_1 same-event S_1 same-event TF_1	NO Sam arrived at eight. He was tired. He rang the bell. NO ?John fell. He was in pain. Mary pushed him. Mary was angry. She pushed John. NO NO I assembled the desk myself. It was beautiful. The drawers only took me ten minutes.
past perf event	just-after S_1 precede S_1 overlap S_1 same-event S_1	Sam had arrived at the house. He rang the bell. Sam arrived at the house. He had lost the key. He rang the bell. NO I had assembled the desk myself. The drawers only took me ten minutes.
past perf activity	just-after S_1 precede S_1 overlap S_1 same-event S_1	Mary had stared at John. He gave her back her slice of pizza. NO NO NO
past perf state	just-after S_1 just-after TF_1 precede S_1 precede TF_1 overlap S_1 overlap TF_1 same-event S_1 same-event TF_1	NO Martha discovered the broken lock. Someone had been in the garage. They rearranged the tools. NO NO Martha discovered the broken lock. Someone had been in the garage. They rearranged the tools. NO NO Mary built the desk herself. She had been happy taking it on. The drawers only took her ten minutes.

For the eventive second sentence of (8b) to be an elaboration of the first sentence, it must occur in a stative form—for example as a progressive (i.e., *She was building a dog house*).

Because of considerations like these, our aim in the implementation work was to treat tense, aspect, cue words and rhetorical relations as mutually constraining, with more specific information such as explicit cue words having higher priority than less specific information such as tense. The main advantage of this approach is that it reduces temporal structure ambiguity without having to rely on detailed world knowledge postulates.

Table 1 lists the possible temporal relations between the eventualities described by two con-

secutive sentences without temporal expressions or cue words, where the first sentence (S_1) may have any tense and aspect and the second sentence (S_2) expresses a simple past event. We constrain S_2 in this way because of lack of space; additional constraints are given in (Hitzeman et al., 1994). For example, if a simple past eventive sentence follows a simple past eventive sentence the second event can be understood to occur just after the first, to precede the first or to refer to the same event as the first (an elaboration relation), but the two events cannot overlap; these constraints are weaker, however, than explicit clues such as cue words to rhetorical relations and temporal expressions. When S_1 expresses a state, it

is possible for the temporal relation to hold between the event described by S_2 and the event or activity most closely preceding S_1 , i.e., the temporal focus of S_1 , here referred to as TF_1 .¹

However, we haven't solved the problem completely at this point: although tense can provide a further constraint on the temporal structure of such discourses, it can also add a further ambiguity. Consider (9):

- (9) Sam rang the bell. He had lost the key.

Clearly, the event described by the past perfect sentence must precede the event described by the first, simple past sentence. However, if a third sentence is added, an ambiguity results. Consider the following possible continuations of (9):

- (10) a. ...Hannah opened the door.
b. ...It fell through a hole in his pocket.

The temporal relation between these continuations and the portion of earlier text they attach to is constrained along the lines sketched before. The problem here is determining which thread in (9) they continue; (10a) continues the thread in which Sam rings the bell, but (10b) continues the thread in which Sam loses the key.

A further ambiguity is that when the third sentence is past perfect, it may be a continuation of a preceding thread or the start of a new thread itself. Consider:

- (11) a. Sam rang the bell. He had lost the key. It had fallen through a hole in his pocket.
b. John got to work late. He had left the house at 8. He had eaten a big breakfast.

In (a) the third sentence continues the thread about losing the key; in (b) the third starts a new thread.²

¹In this chart it appears that whether the tense is simple past or past perfect makes no difference, and that only aspect affects the possible temporal relations between S_1 and S_2 . However, it is important not to ignore tense because other combinations of tense and aspect do show that tense affects which relations are possible, e.g., a simple past stative S_2 cannot have a *precede* relation with any S_1 , while a past perfect stative S_2 can.

²We will not discuss the additional problem that if the final sentence in (11b) is the end of the text, the text is probably ill-formed. This is because

For the problem with multi-sentence discourses, and the "threads" that sentences continue, we use an implementation of temporal centering (Kameyama et al., 1993; Poesio, 1994). This is a technique similar to the type of centering used for nominal anaphora (Sidner, 1983; Grosz et al., 1983). Centering assumes that discourse understanding requires some notion of "aboutness." While nominal centering assumes there is one object that the current discourse is "about," temporal centering assumes that there is one thread that the discourse is currently following, and that, in addition to tense and aspect constraints, there is a preference for a new utterance to continue a thread which has a parallel tense or which is semantically related to it and a preference to continue the current thread rather than switching to another thread. Kameyama *et al.* (1993) confirmed these preferences when testing their ideas on the Brown corpus.

As an example of how the temporal centering preference techniques can reduce ambiguity, recall example (9) and the possible continuations shown in (10). The difficulty in these examples is determining whether the third sentence continues the thread begun by the first or second sentence. For example, in (10a) the preference technique which allows us to choose the first thread over the second is one which assigns a higher rating to a thread whose tense is parallel to that of the new sentence; in this case both *Sam rang the bell* and *Hannah opened the door* are in the simple past tense. In example (10b) the fact that the key is mentioned only in the second sentence of (9) links (10b) with the second thread. To handle an example like (12), we employ a preference for relating a sentence to a thread that has content words that are rated as semantically "close" to that of the sentence:

- (12) Sam rang the bell. He had lost the key.
His keyring broke.

We store semantic patterns between words as a cheap and quick form of world knowledge; these patterns are easier to provide than are the detailed world knowledge postulates required in some other approaches, and result in similar and sometimes more precise temporal

a well-formed text should not leave threads "dangling" or unfinished. This is probably also the reason for the awkwardness of the well-known example *Max poured a cup of coffee. He had entered the room.*

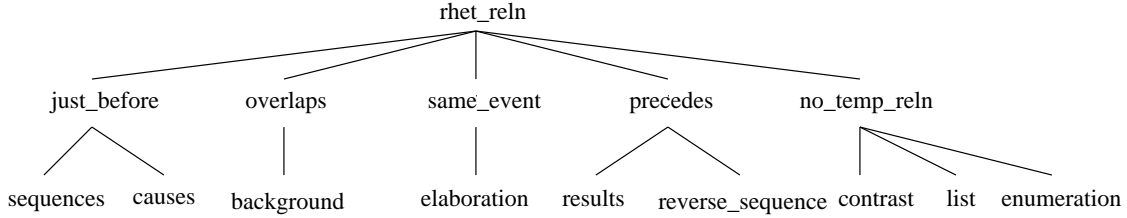


Figure 1: The type hierarchy used for constraints

structures with less processing overhead. Using the semantic patterns we know that *key* and *keyring* are semantically close, and through that semantic link between the second and third sentences we prefer to connect the third sentence to the thread begun by the second.³ The approach to representing semantic relationships we take is one used by Morris & Hirst (1991) wherein the words in the lexicon are associated with each other in a thesaurus-like fashion and given a rating according to how semantically “close” they are. We thus avoid relying on high-level inferences and very specific world knowledge postulates, our goal being to determine the temporal structure as much as possible prior to the application of higher-level inferences.

3 An HPSG implementation of a discourse grammar

Following Scha & Polanyi (1988) and Prüst *et al* (1994), our model of discourse consists of units called Discourse Constituent Units (DCUs) which are related by various temporal and rhetorical relations. A basic DCU represents a sentence (or clause), and complex DCUs are built up from basic and complex DCUs.

In our ALE implementation, a DCU contains the following slots for temporal information:

CUE_WORD: Cues to rhetorical structure, e.g., “because.”

V_AND_NP_LIST: Contains content words found in this DCU, and is used to compare the content words of the current DCU with those in previous threads, in order to rate the semantic “closeness” of the DCU to each thread.

³Semantic closeness ratings won’t help in examples (9) – (10) because there is as strong a relationship between *door* and *bell* as there is between *door* and *key*.

SEM_ASPECT: Contains the semantic aspect (*event*, *state*, *activity*). We have extended the Penn & Carpenter implementation of the HPSG grammar so that semantic aspect is calculated compositionally (and stored here).

RHET_RELN: The relation between this DCU and a previous one. Lexical items and phrases such as cue words (stored in **CUE_WORD**) affect the value of this slot.

TEMP_CENTER: Used for temporal centering; Keeps track of the thread currently being followed (since there is a preference for continuing the current thread) and all the threads that have been constructed so far in the discourse.

FWD_CENTER: Existing threads

BKWD_CENTER: The thread currently being followed

CLOSED_THREADS: Threads no longer available for continuation

TEMP_EXPR_RELNS: Stores the semantic interpretation of temporal expressions associated with this DCU.

TEMP_RELNS: Stores the temporal relations between the eventualities in the discourse.

TEMPFOC: The most recent event in the current thread which a subsequent eventuality may elaborate upon (*same-event*), *overlap*, come *just_after* or *precede*.

TENASP: Keeps track of the tense and syntactic aspect of the DCU (if the DCU is simple).

TENSE: past, pres, fut

ASPECT: simple, perf, prog, perf_prog

To allow the above-mentioned types of information to mutually constrain each other, we employ a hierarchy of rhetorical and temporal relations (illustrated in Figure 1), using the

ALE system in such a way that clues such as tense and cue words work together to reduce the number of possible temporal structures. This approach improves upon earlier work on discourse structure such as (Lascarides and Asher, 1991) and (Kehler, 1994) in reducing the number of possible ambiguities; it is also more precise than the Kamp/Hinrichs/Partee approach in that it takes into account ways in which the apparent defaults can be overridden and differentiates between events and activities, which behave differently in narrative progression.

Tense, aspect, rhetorical relations and temporal expressions affect the value of the RHET_RELN type that expresses the relationship between two DCUs: cue words are lexically marked according to what rhetorical relation they specify, and this relation is passed on to the DCU. Explicit relation markers such as cue words and temporal relations must be consistent and take priority over indicators such as tense and aspect. For example, sentence (13) will be ruled out because the cue phrase *as a result* conflicts with the temporal expression *ten minutes earlier*:

- (13) #Mary pushed John and as a result ten minutes earlier he fell.

On the other hand, if temporal expressions indicate an overlap relation and cue words indicate a background relation as in (14), these contributions are consistent and the RHET_RELN type will contain a *background* value (the more specific value of the two):

- (14) Superman stopped the train just in time. Meanwhile, Jimmy Olsen was in trouble.

4 The algorithm

For reasons of space it is difficult to give examples of the sign-based output of the grammar, or of the ALE rules, so we will restrict ourselves here to a summary of the algorithm and to a very limited rendition of the system output. The algorithm used for calculating the temporal structure of a discourse can be summarised as follows. It consists of two parts, the constraint-based portion and the preference-based portion:

1. The possible temporal/rhetorical relations are constrained.

- (a) If there is a temporal expression, it determines the temporal relationship of the new DCU to the previous ones, and defaults are ignored.
- (b) Lexical items such as cue words influence the value of the RHET_RELN type (See Figure 1).
- (c) If steps (a) and (b) attempt to place conflicting values in the RHET_RELN slot, the parse will fail.
- (d) If there is no temporal expression or cue phrase, tense and semantic aspect also influence the value of the RHET_RELN type (See Table 1), so that rhetorical relations, tense and aspect constrain each other.

2. If more than one possibility exists, semantic preferences are used to choose between the possibilities.

- (a) A “semantic distance” rating between the new DCU and each previous thread is determined. (If there are no existing threads a new thread is started.)
- (b) Other preferences, such as a preference for relating the new DCU to a thread with parallel tense, are employed (See (Kameyama et al., 1993; Poesio, 1994) for details), and the resulting ratings are factored into the rating for each thread.
- (c) If the thread currently being followed is among the highest rated threads, this thread is continued. (This corresponds to temporal centering’s preference to continue the current thread.)
- (d) If not, the DCU may continue any of the highest rated threads, and each of these solutions is generated.

Charts such as Table 1 provide the observations we use to fill in the value of RHET_RELN. Those observations are summarised below. In what follows, the event variable associated with DCU_i is e_i and the TEMPOF of e_1 is the most recent event/activity processed, possibly e_1 itself:

- e_2 can overlap with e_1 if
 - DCU_2 describes a state, or
 - DCU_1 describes a state and DCU_2 describes an activity.

- e_2 can occur just-after the TEMPFOC of e_1 if
 - DCU_2 describes a simple tense event, or
 - DCU_1 describes a complex tense clause and DCU_2 describes a complex tense event, or
 - DCU_1 describes an event and DCU_2 describes an atelic or a simple tense state, or
 - DCU_1 describes a state and DCU_2 describes a simple tense activity.
- e_2 can precede e_1 if
 - DCU_2 describes an event, or
 - DCU_1 doesn't describe an activity and DCU_2 describes a past perfect stative.
- e_2 can elaborate on e_1 if
 - DCU_1 describes an event, or
 - DCU_1 describes an activity and DCU_2 describes an atelic, or
 - DCU_1 and DCU_2 describe states and either DCU_2 describes a simple tense state or DCU_1 describes a complex tense state.

Using this algorithm, we can precisely identify the rhetorical and temporal relations when cue words to rhetorical structure are present, as in (15):

- (15) John fell (e_1) because Mary pushed him (e_2).
 TEMP_RELNS: e_2 precedes e_1

We can also narrow the possibilities when no cue word is present by using constraints based on observations of tense and aspect interactions such as those shown in Table 1. For example, if DCU_1 represents a simple past eventive sentence and DCU_2 a past perfect eventive sentence, then in spite of the lack of rhetorical cues we know that e_2 precedes e_1 , as in (16):

- (16) Sam rang the doorbell (e_1). He had lost the key (e_2).
 TEMP_RELNS: e_2 precedes e_1

Also, when several structures are possible we can narrow the possibilities by using preferences, as in the examples below:

- (17) Sam arrived at the house at eight (e_1).
 He had lost the key (e_2).
 a. ...He rang the bell (e_3).
 TEMP_RELNS: e_2 precedes e_1 ,
 e_3 just-after e_1
 b. ...It fell through a hole in his pocket ($e_{3'}$).
 TEMP_RELNS: e_2 precedes e_1 ,
 $e_{3'}$ just-after e_2

If we allow any of the four possible temporal relations between events, both continuations of sentence (17) would have 17 readings ($4 \times 4 + 1$ reading in which the third sentence begins a new thread). Using constraints, we reduce the number of readings to 4. Using preferences, we reduce that to 2 readings for each continuation. The correct temporal relations are shown in (17).⁴

5 An underspecified representation

By using constraints and preferences, we can considerably reduce the amount of ambiguity in the temporal/rhetorical structure of a discourse. However, explicit cues to rhetorical and temporal relations are not always available, and these cases result in more ambiguity than is desirable when processing large discourses.

Consider, however, that instead of generating all the possible temporal/rhetorical structures, we could use the information available to fill in the most restrictive type possible in the type hierarchy of temporal/rhetorical relations shown in Figure 1. We can then avoid generating the structures until higher-level information can be applied to complete the disambiguation process.

6 Conclusion

We presented a brief description of an algorithm for determining the temporal structure of discourse. The algorithm is part of an HPSG-style discourse grammar implemented in Carpenter's ALE formalism. Its novel features are that it treats tense, aspect, temporal adverbials and rhetorical relations as mutually constraining; it postulates less ambiguity than current

⁴The other reading, in which the third sentence is an elaboration of one of the preceding events, must not be ruled out because there are cases such as *Sam arrived at the house at eight. He rang the bell. He let it ring for two minutes*, in which such elaboration is possible.

temporal structuring algorithms do; and it uses semantic closeness and other preference techniques rather than full-fledged world knowledge postulates to determine preferences over remaining ambiguities. We also recommended using an underspecified representation of temporal/rhetorical structure to avoid generating all solutions until higher-level knowledge can aid in reducing ambiguity.

References

- [Carpenter1993] Bob Carpenter. 1993. *ALE: The Attribute Logic Engine User's Guide*. Laboratory for Computational Linguistics, Philosophy Department, Carnegie Mellon University, version β , May.
- [Grosz et al.1983] Barbara J. Grosz, Aravind Joshi, and Scott Weinstein. 1983. Providing a unified account of definite noun phrases in discourse. In the proceedings of the *21st Annual Meeting of the Association for Computational Linguistics*, pages 44–50.
- [Hinrichs1981] Erhard W. Hinrichs. 1981. Temporale anaphora in englischen. StaatsExamen thesis, Universität Tübingen.
- [Hitzeman et al.1994] Janet Hitzeman, Claire Grover, and Marc Moens. 1994. The implementation of the temporal portion of the discourse grammar. Deliverable D.2.Temporal, LRE 61-062, University of Edinburgh, December.
- [Kameyama et al.1993] Megumi Kameyama, Rebecca Passonneau, and Massimo Poesio. 1993. Temporal centering. In the proceedings of the *31st Annual Meeting of the Association for Computational Linguistics*, pages 70–77, Columbus, OH.
- [Kamp1979] Hans Kamp. 1979. Events, instant and temporal reference. In R. Bauerle, U. Egli, and A. von Stechow, editors, *Semantics from Different Points of View*, pages 376–417, Springer-Verlag.
- [Kehler1994] Andrew Kehler. 1994. Temporal relations: Reference or discourse coherence? In the proceedings of the *32nd Annual Meeting of the Association for Computational Linguistics*, pages 319–321, June.
- [Lascarides and Asher1991] Alex Lascarides and Nicholas Asher. 1991. Discourse relations and defeasible knowledge. In the proceedings of the *29th Annual Meeting of the Association for Computational Linguistics*, pages 55–63, University of California at Berkeley.
- [Morris and Hirst1991] J. Morris and Graeme Hirst. 1991. Lexical cohesion computed by thesaural relations as an indicator of the structure of text. *Computational Linguistics*, 17(1):21–48.
- [Partee1984] Barbara Hall Partee. 1984. Nominal and temporal anaphora. *Linguistics and Philosophy*, 7:243–286.
- [Poesio1994] Massimo Poesio. 1994. *Discourse Interpretation and the Scope of Operators*. Ph.D. thesis, University of Rochester, Department of Computer Science, Rochester, NY.
- [Pollard and Sag1994] Carl Pollard and Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar*. University of Chicago Press and CSLI Publications.
- [Prust1994] Hub Prüst, Remko Scha and Martin van den Berg. 1994. Discourse grammar and verb phrase anaphora. *Linguistics and Philosophy*, 17:261–327.
- [Scha and Polanyi1988] Remko Scha and Livia Polanyi. 1988. An augmented context free grammar for discourse. In *Proceedings of the 12th Conference on Computational Linguistics*, pages 573–577, Prague, August.
- [Sidner1983] Candace L. Sidner. 1983. Focusing in the comprehension of definite anaphora. In M. Brady and R. Berwick, editors, *Computational Models of Discourse*. MIT Press, Cambridge, MA.